DOCUMENT RESUME

ED 381 757 CS 012 113

AUTHOR Stephens, Diane; And Others
TITLE Assessment and Decision Making in Schools: A Cross-Site Analysis. Technical Report No. 614.
INSTITUTION Center for the Study of Reading, Urbana, IL.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Apr 95
NOTE 33p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Administrator Behavior; Case Studies; *Decision Making; Educational Practices; Elementary Education; Evaluation Criteria; Evaluation Research; *Institutional Characteristics; *Teacher Administrator Relationship; *Teacher Behavior
IDENTIFIERS School Culture; *Teaching to the Test

ABSTRACT

Using a case-study-approach, a study sought to describe what assessment looked like in four school districts (two schools per district, two classrooms per school). Interviews were conducted with students, parents, teachers, principals, and central office staff to understand assessment from multiple perspectives. Teachers were interviewed prior to and after three half-days of observation to understand assessment as part of classroom practice. Results indicated that the meanings of particular concepts, such as assessment, curriculum, and accountability, varied significantly across districts. The salient relationship was not the one between assessment and instruction, but rather the relationship of each of these to the decision-making model of the district. Generally, when assessment-as-test did appear to drive instruction, this relationship seemed to be an artifact of a model in which individuals ceded authority for decision making to outsiders. When assessment-as-test did not appear to drive instruction, this relationship seemed to represent a model in which individuals maintained the authority to make decisions within the framework of their individual and collective philosophies. Findings suggest that assessment-as-test does not necessarily drive instruction, and that when assessment-as-test does drive instruction, it does not drive it in a way that might be considered good instruction. (Contains 48 references and a table of data. The interview questions, and the observation and interview coding systems are attached.)
(Author/RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 381 757

Technical Report No. 614

ASSESSMENT AND DECISION MAKING IN SCHOOLS: A CROSS-SITE ANALYSIS

Diane Stephens
University of Hawaii, Manoa
P. David Pearson
University of Illinois at Urbana-Champaign
Colleen Gilrane
University of Tennessee-Knoxville
Mary Roe
University of Delaware
Anne Stallman
Metritech Corporation
Judy Shelton
University of Wisconsin, Milwaukee
Janelle Weinzierl
University of Illinois at Urbana-Champaign
Alicia Rodriguez
University of Illinois at Urbana-Champaign
Michelle Commeyras
University of Georgia

April 1995

Center for the Study of Reading

TECHNICAL REPORTS

College of Education
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
174 Children's Research Center
51 Gerty Drive
Champaign, Illinois 61820

CS012113

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

F. Fisher

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CENTER FOR THE STUDY OF READING

Technical Report No. 614

ASSESSMENT AND DECISION MAKING IN SCHOOLS: A CROSS-SITE ANALYSIS

Diane Stephens
University of Hawaii, Manoa
P. David Pearson
University of Illinois at Urbana-Champaign
Colleen Gilrane
University of Tennessee-Knoxville
Mary Roe
University of Delaware
Anne Stallman
Metritech Corporation
Judy Shelton
University of Wisconsin, Milwaukee
Janelle Weinzierl
University of Illinois at Urbana-Champaign
Alicia Rodriguez
University of Illinois at Urbana-Champaign
Michelle Commeyras
University of Georgia

April 1995

College of Education
University of Illinois at Urbana-Champaign
174 Children's Research Center
51 Gerty Drive
Champaign, Illinois 61820

MANAGING EDITOR
Technical Reports
Fran Lehr

MANUSCRIPT PRODUCTION ASSISTANT
Delores Plowman

Abstract

Using a case-study approach, a study sought to describe what assessment looked like in particular classrooms of particular schools located in four particular districts (Alpha, Beta, Gamma, and Delta). Interviews were conducted with students, parents, teachers, principals, and central office staff to understand assessment from multiple perspectives. Teachers were interviewed prior to and after 3 half days of observation to understand assessment as part of classroom practice. The study found that the meanings of particular concepts, such as assessment, curriculum, and accountability, varied significantly across districts. The salient relationship was not the one between assessment and instruction, but rather the relationship of each of these to the decision-making model of the district. Generally, when assessment-as-test did appear to drive instruction, this relationship seemed to be an artifact of a model in which individuals ceded authority for decision making to outsiders. When assessment-as-test did not appear to drive instruction, this relationship seemed to represent a model in which individuals maintained the authority to make decisions within the framework of their individual and collective philosophies. The study revealed that assessment-as-test does not necessarily drive instruction, and that when assessment-as-test does drive instruction, it does not drive it in a way that might be considered good instruction.

ASSESSMENT AND DECISION MAKING IN SCHOOLS: A CROSS-SITE ANALYSIS

When we began our study of the relationship between assessment and instruction in 1988, our goal was simple: We had heard stories and read reports, both negative and positive, about the relationship between assessment and instruction, and we wanted to understand whether, to what degree, or in what sense the oft-cited assertion that "assessment drives instruction" accurately characterized what was happening in classrooms, schools, and districts.

Assessment, Instruction, and Curriculum

We began our study by reviewing the literature. We found that when educators discussed the relationship between assessment and instruction, assessment was almost exclusively defined as tests developed outside the classroom. Based on this definition, educators offered two opinions about this relationship. The first is that assessment does not drive instruction. This position (Haney, 1984; 1985) argues that testing has affected only surface characteristics of instruction and that the decisions teachers make on a daily basis are not affected by externally imposed testing. In short, as Haney argued, all of the fuss about the tyranny of the test is little more than scholarly rhetoric with no basis in the reality of everyday schooling.

The second position is that assessment does drive instruction. Individuals who take this stance (Brandt, 1978; Brookover, 1987; Burry, 1981; Cohen, 1988; Johnston, 1987; Madaus, 1985; Popham, Cruse, Rankin, Sandifer, & Williams, 1985; Stedman, 1987) argue that what is tested determines what is taught. These scholars differ, however, on the question of whether assessment should drive instruction. Individuals taking the pro position (Brookover, 1987; Cohen, 1988; Popham et al) argue that assessment is a viable, even responsible, means of controlling what happens in classrooms. Individuals taking the con position (Brandt, 1978; Burry, 1981; Calfee, 1987; Johnston, 1987; Valencia & Pearson, 1987) argue that when teachers teach to a test, the curriculum is narrowed, and teachers and students are robbed of their curricular birthright to determine what happens in classrooms.

In the half-decade that has ensued since our study began, the argument against allowing tests to control curriculum and instruction has gathered momentum gradually but consistently. The curriculum-narrowing phenomenon has been studied in detail (Koretz, Linn, Dunbar & Shepard, 1991; Smith, 1991), and we have learned more about the disempowering effect of tests on teachers' sense of professional efficacy (Smith, 1991).

Perhaps the most dramatic demonstration of the curriculum-narrowing impact of teaching to the test occured in a study by Koretz, Linn, Dunbar, and Shepard (1991). Working with a large district that had adopted a high-stakes disposition to assessment (everyone's scores—kids, teachers, and schools—are a matter of public record, and there are consequences to low scores), Koretz and his colleagues (1991) examined the subtle effects of teaching to the test over a five-year period. In 1987, the district switched from one popular standardized test to a second, and their average school mean dropped in comparison to the previous year, over a half grade level. In 1988, 1989, and 1990, the average scores, again computed at the school level, rose substantially each year, and by 1990, for example, their third graders were almost a grade higher than they were in 1987, the first year of the new test. To evaluate the subtle effects of teaching to the test, in 1990 they readministered the standardized test that they had dropped in 1987. Compared to the 1990 administration of the new test, their scores dropped off a half grade level. To counter the argument that the between test differences could be accounted for by differences in norming populations they also compared Test A in 1990 to Test A in 1986, finding that average school scores had dropped off a full grade level. About the only conclusion that can be drawn is that the growth from 1987 to 1990 in performance on standardized Test B was due to teaching to the test.

They made one other comparison. In 1990, they gave alternative tests (what we now call authentic assessment or performance assessment measures) in both mathematics and reading. In general, these were more like everyday classroom assignments (solving math problems on your own, writing in response to reading). They also administered these alternative tests and the "new" standardized test in a district that, while demographically similar, had avoided high-stakes frenzy. Then they compared the two districts performance on the two types of assessment. Regardless of whether one looked at overall test scores, subtest scores, or specific item types, the pattern was consistent: The high-stakes district looked just as good as the other district on the standardized test but scored consistently lower on the alternative assessments. Again, the conclusion that teaching to the test narrows the curriculum is hard to avoid.

The influence of tests in shaping school curriculum has created great concern among many educators (Shepard, 1989). For example, Smith (1991) found that teachers are very sensitive to the publication of test scores. They are willing to alter their curriculum to avoid low test scores on a test they do not believe in, even though the practices they engage in to raise test scores result in personal feelings of "dissonance and alienation" and "guilt" about the harm they feel they are inflicting on children.

Ironically, teaching-to-tests essentially renders most of them invalid (Haladyna, Nolen, & Hass, 1992). Teaching to a test can lead to test-score pollution, a phenomenon that occurs when there is a rise or fall in measured performance without a concomitant change in the underlying construct that is allegedly measured by the test. Multiple-choice tests are typically built on the assumption that no one ever teaches to them directly; hence, they can serve as perfectly reasonable "barometers" of achievement for some construct. But their measurement qualities crumble when the tests are required, either intentional or incidentally, to serve as a blueprint for a curriculum.

In reviewing research concerned with assessment for students of diversity, García and Pearson (1994) have found that the problem of undue curricular influence is even more severe for low-income students. Teachers of low-income students tend to be held more accountable (or at least they feel that they are more accountable) to tests (Center for the Study of Testing, Evaluation, and Educational Policy, 1992; Dorr-Bremme & Herman, 1986; Madaus in Rothman, 1992).

Herman and Golan (n.d.) found that teachers in Chapter 1 classrooms reported "more emphasis on testing, less school attention to broader instructional renewal, more adjustments made to instrucational planning to incorporate aspects of the test, more classroom time spent on test preparation activities, and less classroom time spent on non-tested subjects and skills" (p. 2). If one makes the reasonable assumption that most of these tests feature discrete skills, then it is not difficult to understand the accusation (García, Pearson, & Jiménez, 1994) that low-income students are much more likely than other students to receive a fragmented, skills-based curriculum.

This phenomenon is not limited to literacy assessment. A survey of 2,200 mathematics and science teachers, augmented by intensive visits to six urban sites, revealed that teachers of low-income students were the highly likely to teach to a test (Center for the Study of Testing, Evaluation, & Educational Policy, 1992; Rothman, 1992). One of the teachers in the study, a fifth-grade teacher in an inner city school, explained that she had been using the mathematics curriculum guide "to identify objectives in order to teach to the test" because a certain percentage of students in each district school had to attain a cut-off score on the standardized test or the district would be taken over by the state (Center for the Study of Testing, Evaluation, & Educational Policy, 1992, volume 1). The consequence was that she had "little time to bring in things—connect things" (p. 1). So she followed the textbook 95% of the time because it matched the items covered on the state-required standardized test. Sadly, an analysis of the test, along with five other popular standardized achievement tests in mathematics and science (grades 4, 8, and high school) as well as sample textbook texts, "indicated that only a handful of the questions measured the kinds of conceptual knowledge and problem-solving abilities reformers (e.g., NCTM, 1989)

say should be integral to instruction in those fields" (Madaus and his colleagues as cited in Rothman, 1992, p. 1).

Over the past few years diverse solutions have been offered to these dilemmas. Admitting the futility of trying to shelter curriculum and instruction from the authority of tests, many have become advocates for better tests. Suggestions for improving tests include finding alternatives to pen-and-paper tests (Guthrie & Lissitz, 1985), developing formative testing schemes (Brandt, 1978; Conner et al, 1985), and developing tests that help us understand why particular overall outcomes were achieved (Cohen, 1988). A collaboration of teachers, researchers, and policy makers in Michigan (Wixson, Peters, Weber, & Roeber, 1987) built the state assessment upon the best information and perspective that could be gathered from advances in reading theory, research, and practice. Similarly, Valencia and Pearson (1988) suggested that assessment reforms should focus on orchestrating, rather than isolating, skills and that new tests needed to move beyond the tyranny of the text and acknowledge all relevant factors—the reader, the text, and the context—in a more balanced fashion. Arguing vehemently for assessment reform, leaders in the New Standards Project (e.g., Simmons & Resnick, 1993) have argued that because assessment drives instruction and since misguided assessments have driven us into the curricular swamp in which we are currently mired, new, more virtuous tests (which they define as a combination of performance assessments and portfolios) can lead us to higher, more thoughtful, and more empowering curricular ground.

Others have argued for either eliminating tests or narrowing their range of influence (Brandt, 1978; Burry, 1981; Calfee, 1987; Stedman, 1987). Those who advocate this shift also tend to favor alternative forms of assessment—such as portfolios (Berlak, 1978; Johnston, 1987; Valencia, McGinley, & Pearson 1990) or school-site evaluation teams (Berlak, 1978).

Interestingly, then, there are advocates for alternative assessments on both sides the assessment-instruction dilemma. Except for those who want to use assessment as the wedge for curriculum reform (e.g., Simmons & Resnick, 1993), what most of the advocates for alternative assessment share is a common commitment to the view that the authority for assessment ought to originate in classroom rather than the boardroom or the statehouse. They have adopted a more situated and constructivist view of assessment (Johnston, 1987; Tierney, Carter, & Desai, 1991). While some (e.g., Tierney et al, 1991; García & Pearson, 1994) recognize the possibility that data originating in classrooms may ultimately find its way into the accountability and policy milieux, the more general position is that assessment is more properly used as a tool for making decisions about individuals or classroom programs rather than school-, district-, or statewide programs (e.g., Hancock, Turbill, & Cambourne, 1994). Some advocates argue strongly for connecting it to curriculum and instruction, but unlike those who capitalize on this connection for shaping policy (e.g., Simmons & Resnick, 1993), these advocates situate the connection within a classroom context in which teachers and students determine its role (Tierney et al, 1991). Thus, assessment does not drive instruction but follows naturally from particular conceptualizations of curriculum and teaching. Starting from the same constructivist perspective, others emphasize the role of the child in the process (e.g., Hansen, 1994); they cite the importance of placing responsibility for assessment in the hands of those most affected by it (teachers and students) and making sure that students are involved in every stage the process, from determining what will be assessessed, when and how it will be assessed, and, most important, how it will be interpreted (Hansen, 1992; 1994; Graves & Sunstein, 1992; Tierney, et al, 1991). Not surprisingly those who take this constructivist view also emphasize professional development as a means of helping teachers use assessment to become better decision makers (Calderhead, 1988; García & Pearson, 1994; Guthrie & Lissitz, 1985; Johnston, 1987; Tierney & McGinley 1988).

Purpose

While the literature gave us, and continues to give us, good reason to be concerned about the negative curricular and professional impact of tests, we have found little insight concerning what all of these policy considerations mean for daily life in classrooms. In contrast to the sweeping policy perspectives that we found, we wanted to understand how standardized tests and other assessment tools impacted lives in particular classrooms in particular districts. What was classroom life like in a school that was attempting to raise its test scores? Were daily patterns different in those schools from the patterns that characterized the lives of teachers and students in schools which were not highly invested in raising test scores? We had read of teachers who "taught to the test;" we wanted a closer and yet broader understanding of what that meant. We wondered, for example, about the relationship of textbook orders (kind and company) to the test. Might an individual feel unaffected by test pressures and yet be required to use materials that had been specifically chosen to match a particular test or even test items? And what about policies for passing versus retention? Might a teacher feel relatively free from test pressures during the year but then be told that only students who achieved certain reading levels could pass to the next grade, a grade in which standardized tests were admistered?

To move our understanding from the abstract (research says that testing drives or does not drive instruction) to the concrete (what does this mean in the lives of particular teachers/schools/districts?), we decided to conduct case study research. Aware that the impact of assessment on instruction could be subtle (the impact might emerge in a decidion to adopt a particular text that they believed would help raise tests scores), we situated our study within the context of school and classroom decision making.

Method

We chose to work within a qualitative research paradigm because, with its social constructivist orientation, we felt it was more consistent with and sympathetic to our interest in the meaning that participants made of assessment and the relationship between assessment and instruction. Of the various qualitative options, a case-study approach seemed particularly well suited to our needs. First, a case study examines a specific phenomenon, such as, in our case, the assessment-instruction link. Second, a case study can illustrate the complexities of a cultural event (Hoaglin, 1982). For example, assessment, a cultural activity, occurs in a cultural location, a school, and is conducted by members of that culture, teachers and students. A case study allows the exploration of this interplay. Finally, case-study methods, with their emphasis on the human instrument, allow a thorough exploration of information, an integration of information across sources, and response to serendipity (Lincoln & Guba, 1985). The collection of observational data and the conduct of interviews, our means of data collection, followed guidelines typical of qualitative inquiry (Glesne & Peshkin, 1992).

Site Selection

Functioning under real-life constraints of time, money, and personnel, we limited the study to four districts, two schools per district, two teachers per school. We selected districts that we thought would offer different approaches to decision making. Once these decisions had been made, we contacted central office staff in four districts, explaining the study and asking if their district would be willing to participate. The four districts were subsequently given the pseudonyms Alpha, Beta, Gamma, and Delta.

In our conversations with school personnel, we explained that our interest was in the relationship between standardized tests and instruction. We told them that we wanted to situate both tests and instruction within a broader framework of instructional decision making so that we could better understand the more subtle influences of one on the other, (e.g., textbook purchasing policies). We also explained that we were interested in the seldom discussed assessment-that-was-not test (e.g., teacher observation or informal diagnostic procedures) and the relationship of that form of assessment to

instruction. All participants therefore understood that we were interested in decision making as it related to assessment (both as test, and not-as-test) and instruction.

Alpha, located in a midwestern university town of 40,000 people, was selected because we had reason to believe that its teachers had a great deal of autonomy; we believed it might "anchor one end" of a decision-making continuum. Thinking it might present a more "top-down" administrative perspective, we selected Beta, which was located in a midwestern city of about 60,000 containing a university, community college, and some manufacturing and service industries. We believed that Gamma, located in the suburbs of a major midwestern city, would allow us to examine the role of assessment in a district in which high levels of student performance were expected and achieved. Delta, in the same location, heard about the study and asked to be included. They were concerned about what they perceived to be their low reading scores and hoped that participating in the study would help them better understand their reading program as well as what they might do about their scores.

Demographic data about these districts are provided in Table 1. It is interesting to note that the two downstate districts are considerably more ethnically diverse than the two suburban districts, although Delta is itself much more diverse, particularly linguistically, than Gamma. These pre-existing demographic differences presented us with opportunities to learn whether they were associated in any systematic way with differences in assessment practices. The data for Gamma clearly corroborate the "image" that motivated us to work with the district in the first place; it is clearly a high-profile, ethnically homogenous, high-achieving district. The variability in state reading scores is also interesting in view of the fact that Delta, the second suburban district, was most concerned about test scores, when, in fact, its scores were higher than those obtained by either Alpha or Beta. In our initial discussions, we heard little from personnel in either Alpha or Beta about their low state test scores. The discrepancy between actual scores and perceived problems provided us with a unique opportunity to study the influence of community expectations on assessment practices.

[Insert Table 1 about here.]

The districts responded differently to our expressions of interest. In Alpha, central office staff notified all teachers that we wanted to conduct a study and asked them to contact us if they were interested. In that district, seven teachers in one building participated and two in another. In Beta, central office staff decided which buildings and teachers would articipate. In Gamma and Delta, central office staff invited teachers and principals to a meeting to hear about the study and then chose two schools from among those interested.

Data-Collection Techniques

Observations. One member of the research team observed each teacher's classroom on three occasions. The expanded fieldnotes from these observations were shared with teachers for their feedback. Their comments became part of the data base. These observations provided specific documentation of classroom events and of assessment and decision making within them.

Interviews. To understand assessment, instruction, and decision making from multiple perspectives, we interviewed students, parents, teachers, principals, and central office staff. Each participant, with the exception of teachers, was interviewed once. For each category of respondents, questions written in advance (see Appendix A) provided a general direction and consistency across interviews. The interviews, however, remained open ended. Only rarely were all the probes used; indeed, often only the first question was asked as written. Then, in the conversation that ensued, researchers asked contextually embedded questions that followed the lead of the person being interviewed. In all cases, we kept track of the content of each interview as it was occurring to make certain that all questions had been answered even if they had not been explictly asked. In this way, we hoped to understand how the

participants approached assessment, instruction, and decision making. This understanding, we reasoned, would help us build a framework that encompassed answers to original questions as well as emerging questions, answers, issues, and perspectives.

To understand assessment as a part of classroom practice, we also interviewed teachers prior to and after each of three observations lasting one half day. In contrast to the semi-structured nature of the initial interview, when we talked with teachers, we asked questions that allowed us to understand what we might expect to see and what had occurred as well as to understand the reasons behind the actions we had observed. For example, in one case, after clarifying that there were indeed three reading groups, and that there had been three since the beginning of the year, we asked the teacher general questions such as how she had decided on three groups, who should be in each group, and what materials to use. We also asked specific questions tied directly to the observation, for example, "I noticed that you went around the group today, making sure that everyone had their finger on a particular place in the text, could you please talk to me about the reason you did that?" With the participant's permission, interviews were tape recorded and transcribed. Copies of interviews were returned to the teachers, principals, and central office staff for their comments. Their feedback became part of the data base used in our analyses.

Data Analysis

The amount of data collected across all four sites was extensive. When fieldnotes were elaborated and audio tapes transcribed, we had more than 3,000 pages of text. Each interview, observation, and response was entered into a qualitative data base and coded descriptively to facilitate analysis. The codes were determined fairly early in the data-collection process and primarily described categories that we felt it might be helpful to look at more closely later. Subsequently we made little direct use of the codes, relying instead on more traditional methods of qualitative data analysis (e.g., reading and rereading all of the data to identify patterns, conductive negative case analyses to evaluate competing hypotheses, and identifying linkages across distant entries). We found that the codes often hindered rather than supported our identification of patterns because, when we retrieved information by code, the information was stripped of the surrounding context. We did, however, use the codes to look systematically at particular categories of remarks and/or to investigate patterns we thought might be particularly salient (e.g., the proportion of teacher talk about standardized versus informal assessment). assessment. Copies of those coding systems are provided in Appendix B.

Analysis of cases. The analyses of the first three districts were conducted concurrently, with one member of the research team taking primary responsibility for one site. A constant comparative approach was used in the analysis. Each researcher read and re-read the data, looking for and identifying patterns in the data. Once patterns had been identified, the data were read at least one more time to look for evidence that might disconfirm those patterns. The researcher then detailed those patterns in a case study that aptly captured what we had learned about assessment and instruction in that district. Meanwhile, members of the research team continued to meet with each other, sharing possibilities and patterns.

Once we had preliminary drafts of these case studies, we began, as a group, to analyze data from the fourth site. When that analysis was completed, we substituted pseudonyms for the names of all participants and sent each district a copy of its case study and asked for feedback. Based on responses, we made changes, as appropriate, in the case studies and then sent copies of all four case studies to each participant. All four case studies were then published as Center for the Study of Reading Technical Reports (Rodriquez et al., 1993; Shelton et al., 1993; Stephens et al., 1993; Weinzierl et al., 1993).

Analysis across sites. Following the completion of each case study, we returned to the interviews and observations and began to analyze data across sites. After we had generated new categories that we all

believed were helping in explaining the patterns we saw across sites, the data were read and coded one more time. We then reread to search for negative cases, instances that might prompt us to reconsider our analysis of themes and patterns.

RESULTS AND DISCUSSION

Patterns Within the Districts

While our ultimate goal was to examine patterns across the four sites, we wanted to ensure that our cross-site analysis was grounded in an in-depth examination of the data from each school district. We therefore began by exploring each of the four districts as separate cases, each with its own culture and integrity (for detailed reports of each district, see Rodriguez et al, 1993; Shelton et al., 1993; Stephens et al., 1993; Weinzierl et al., 1993). We drew maps, made lists of committees, and designed flow charts to trace decision making through the organizational patterns of the districts. In essence we created a sketch of each district by pulling its unique and particular characterstics into the foreground. We developed narrative accounts, metaphors, key words, and visual models to portray the particularity of each district. From these separate analyses, we developed a sense of how decisions were reportedly made in each district, who controlled what, and how materials were chosen. We knew what tests were given, by whom, and how those results were used. We knew too about the other kinds of assessment that occurred in each district and how those data were used and valued.

As of result of these analyses, we came to understand that the relationship between assessment and instruction could not be compared across districts without first discussing what those terms meant within each district. In retrospect, we realized that when we began the study, we had assumed that there was a homogeneous culture called "school," and that we would study the relationship between assessment and instruction in four districts within that culture. What we came to understand is that we instead studied at least four different cultures. The meaning of particular concepts—assessment, curriculum, accountability—varied so significantly across districts that to "do school" in one district was not the same as to "do school" in another. The sketches of each district provided in the following sections are based upon these individual case studies. They highlight each district's ideas about curriculum, accountability/responsibility, and assessment.

Alpha

In Alpha, a district-wide Curriculum Council oversaw all curriculum writing. The council consisted of teachers and administrators representing each school in the district. When the council determined that a particular curricular area needed attention, they advertised for teachers to chair a curriculum committee. The position carried with it a $2,000 stipend. Teachers applied and were interviewed by the Council as candidates for this position. The teacher who was chosen, in turn, appointed committee members from among a set of volunteers.

What emerged from these committees was a broad-based, very general curriculum, more than likely a set of goals or standards, as explained by the superintendent:

> it's more of a philosophy than a set of things to teach. Our district curriculum does not produce courses of study. Teachers, together or alone, produce courses of study consistent with the district curriculum . . . . So it makes sense to talk of a school curriculum or even a classroom curriculum . . . . We might identify areas, goals, and even choices of materials, but we never identify any particular set of materials as our curriculum.

These committees were not, as several Alpha participants pointed out, textbook committees: "They don't adopt materials; they write curriculum." As one of the Alpha teachers, explained:

> The parameters or framework is that the district has a curriculum guide and within that guide you're free to choose what is applicable for your grade level . . . how the lesson is taught or what materials you should use would be completely up to you . . . . As a matter of fact there's no one text, no one thing for anything that we teach . . . .

Curriculum in Alpha was a construction of teachers and students within classrooms, guided by the philosophy of the district, and, within that, by the philosophy of the individual teacher.

In Alpha, accountability was sometimes contrasted with responsibility. As one administrator noted, "I think what accountability does is to focus you on the entire group, whereas responsibility focuses you on the individual kid." Paradoxically, because of the district emphasis on individual autonomy, accountability came to mean responsibility. This theme was subtle, and yet woven into almost every conversation about assessment and instruction. As one principal noted:

> I would expect the teachers to be fully informed about their students abilities, needs, and capabilities in order to make fully informed decisions. I want a knowledgeable person in that position. I expect the person to be able to handle all of that and we would explore all of that whenever we interview a prospective teacher.

Teachers shared this sentiment:

> If [a child] came into my room reading at a third grade level, I would expect him to leave reading at a fourth grade level. If [a] child comes in knowing some letters, my expectation would be that at the end they would be doing some reading. You certainly don't have a certain level and if everybody makes it there, then we're fine. I think everyone has to be called upon . . . to dig down deep and move from there and beyond. I think we have to accept where they come in at and say to them, okay, here's where you're at, let's see if we can get over here to this point.

Accountability, as responsibility, meant knowing each child well so that instructional decisions could be grounded in knowledge of the individual and of the individual within the group. Qualitatively, a "good" job, relative to accountability, would mean that each teacher could paint a portrait of each child. Teachers often kept folders and anecdotal records for each child so that they would have documents available to show themselves, the child, and others how progress was constituted.

In Alpha, we realized that if we were to do a collage that represented our sense of the district, we would include the words *autonomy, individual, responsibility, choice*, and *professionalism*. Neither the word assessment nor the word test would be a part of the collage. Rather than exist as a separate category, assessment was woven into Alpha's concepts of both accountability and curriculum. Our collage would also contain portraits of individuals. Earlier, we noted that accountability in Alpha meant being able to paint portraits of individual children. Interestingly, in our interactions with Alpha participants, teachers also painted portraits of themselves and each other as unique individuals. Indeed, the individual images were so strong that they did not seem to be contained in any structure.

For a visual metaphor, we chose a canvas to represent Alpha. It captures our sense of Alpha as a community focused on attention to the individual—to the individual teacher as professional, to the individual student as learner. Their motto seemed to be "Nurture Each Individual."

Gamma

Gamma calls to mind very different words, phrases, and images. When we think of Gamma, we think not of individuals, but of teams. We envision team members supporting and encouraging each others' efforts to do their very best. What was particularly fascinating about Gamma was the number of teams with which each teacher was involved. Our analysis of the data suggested that each Gamma teacher belonged to at least two teams other than the grade-level team. Curriculum, as group-supported teaching strategies and instructional materials, was one of the issues discussed in team meetings. Curricular suggestions were formalized in cross-school team meetings, and workshops then served to share these ideas with other teachers. As one teacher explained:

> This district is really big on inservice type training. And then those who are trained come back and help other people in the building. It's kind of a feeder system. You know you can do this, now you can feed it to other people. It works really well. Especially if you're in a place where you really value your colleagues' opinions and you value how they teach and what they do.

Teachers volunteered to pilot new materials and strategies in their classrooms. They then observed each others' classrooms, discussed the innovation and, if the idea were considered a success, curriculum became a district initiative in the form of, "This is what we'd like for you to do." Teachers had the option of accepting, rejecting or modifying curriculum. Once curriculum was in place, surveys were sent home to parents for additional input.

In Gamma, teams formed intimate units that insured the flow of ideas and encouraged communication. Gamma educators explained that administrators and teacher leaders from Gamma had been trained in a particular model of collegial decision making. One of the teachers selected for leadership training talked about the influence of this training:

> [When we came back] we couldn't go to the teachers and say, "Okay, this is what we learned and this is what you should do." It was supposed to be teacher initiated and that takes a whole lot more thought, and working with kid gloves than saying, "Okay, this is what we are going to do." Not only was it to be teacher initiated as to what we were going to do, it was going to be teacher initiated as to how it was going to be implemented.

Our model of decision making in Gamma is very complex because of all the consituencies involved; the model has to capture both how the organizational patterns allowed for various grouping arrangements, all of which kept lines of communication open; and that the community was a part of the dialogic process. In Gamma, assessment was seldom singled out as a separate topic, instead it became one topic of conversation within multiple, on-going dialogues. Gamma wanted to be the best, and assessment, defined as data from tests and data from teachers, was one means of achieving that goal. Accountability was two-fold, including both a desire to achieve the district motto of being the best and a sense of responsibility to colleagues, students, and parents.

Many of Gamma's ideas about organization and communication had parallels in the corporate world. Therefore, we chose a corporation to symbolize school in Gamma and decided that "Be the Best" was an apt motto.

Delta

In both Alpha and Gamma, one dominant, clearly articulated decision-making structure emerged. For our metaphorical model, we drew Alpha as a canvas and Gamma as a corporation. Representing Delta

as one visual image was harder for, in many ways, Delta seemed to us to be comprised of many images—a district in transition, a district undergoing a metamorphosis. In our discussions with Delta educators, for example, we heard about three approaches to decision making and of the tension that resulted when the approach anticipated was not the approach used. We also found evidence that the tension and debate were generative, that they were part of a change process. As the superintendent explained,

> We are beginning to understand what we don't know . . . I can see everybody being ready for change and that's the first major step . . . I think we're ready. If enough people can come with information we can use, I think my faculty is ready.

At the time we were there, one administrator described the decision-making process as three-pronged: (a) a top-down approach, in which decisions were made by administrators and "shared with the staff," (b) a more "democratic" model in which "input was sought from the staff," and (c) a committee structure in which teachers and administrators worked together.

The tensions among these models can perhaps best be understood by examining the textbook adoption process. In Delta, the textbook was the curriculum. As one teacher explained, the charge to the curriculum committee was to rewrite the objectives to match the basal series currently being used. Textbook adoption occurred every six years. A curriculum committee selected two or three textbooks for adoption, and teachers voted to indicate their preferences. The last reading textbook adoption had proceeded in this fashion: Textbook publishers made presentations to central office personnel. Three series were selected and subsequently approved by various committees. Teachers then voted. The majority of teachers voted for either series A or B. Because there was no consensus, a decision was made by the school board to adopt series C. Teachers had expected the textbook decision to be made in the "more democratic model" and expressed anger about the "top-down" manner in which the decision was made.

Putting this decision in historical perspective, the superintendent explained that, at the time, he had thought the district "had too many things going," and that they "needed to structure things a little more." These concerns led to adopting a single set of materials for reading. At the time of our visit, however, he noted that the "reading committee was coming back and saying that maybe this is a little too structured." Indeed, because the formation of the most recent reading committee, the superintendent saw his role as that of a "rubber stamp":

> I've got a reading committee who are experts compared to me and I just have to listen to them and hope I am intelligent enough to make the right decision, that is, to support them or not. So I think our decisions are made where they need to be, and that's with a group of teachers who have the interest or the knowledge of that area, have spent a considerable amount of time working on it and then come through the system with their recommendations.

In Delta, we found several other examples of issues that contributed to a sense of tension in the district. A retention committee had been formed, but it was composed solely of administrators. A pacing policy written by that committee, not the judgment of individual teachers, determined when a child could move to the next book in a series and whether or not the child would be retained. Recess had been eliminated and lunch shortened in order to increase the amount of time devoted to instruction.

State test results were yet another source of tension. Educators in the district felt embarrassed by the scores their students received on the state reading test and were actively seeking ways to improve their reading program. Just as Delta teachers and administrators felt accountable to the state and to the public for their standardized test scores, teachers felt accountable to principals and principals felt accountable to central office administrators. Principals, for example, were asked to monitor the pace

of instruction by collecting basal end-of-unit test scores in math and reading from each teacher every Friday. This information was entered into a computer data base and a record of scores was sent to the assistant superintendent and the superintendent. This policy was designed to ensure that all teachers in the district covered math and reading at the same pace.

Delta educators spoke openly of the tensions within their district and offered diverse solutions. One principal, for example, noted that he was not sure how much teachers should be involved in decision making. As he explained,

> I think teachers basically want to teach. They want to go into their classrooms and be responsible for what they're teaching their children. As far as being involved in a lot of committees that share in decision making and various things, I don't think they really want to be out of their classroom. I think they'd rather be in their classroom and have someone else make the decisions for them. This might be kind of pompous but I'm not even sure they're prepared to make those decisions. I'm not sure they have sufficient background experience to make them.

Other educators in the district held similarly strong, although contrasting, ideas about how decisions should be made. One teacher, for example, described a meeting in which she was told that according to the retention policy, she could only recommend that a child repeat a grade if the child had failed the first quarter of the year:

> I said, what policy? And that is the first I had heard of the policy. So things like that, I think, teachers need to be involved in. Now when I spoke with the superintendent on it, he agreed with me. He did agree. He said the teachers will be given a chance to react to the policy . . . . Not that that [will] change it.

Another teacher explained how she felt about the role of teachers in the decision making process:

> It seems like lately the picture is that there are a lot of decisions being made at the top and that there are a lot of committees being formed that teachers are not on...That's very frustrating to me . . . . [For example] we wanted phonics books in the fall, we were out of money. Talk about decision-making, we want them, we need them, we consider them important, it's not our only tool, we are using it as one of the tools and we are getting better results . . . . They just don't seem to listen.

Assessment, instruction, accountability, and responsibility were all parts of this complex and highly charged school climate. Predominantly, but not exclusively, assessment meant tests, and educators at all levels felt accountable to the public and the state for their test scores. Instruction was predominantly, but not exclusively, a matter of covering the materials in the district adopted books and ensuring that students had passed mastery tests at a level of 80% accuracy before progressing to the next book in the series. Responsibility was predominantly, but not exclusively, to the tests and to the material.

In deciding how to portray our impressions of Delta in a sketch and with a few key words, we decided that the most apt visual metaphor to capture the sense of ambiguity and transition was a schoolhouse in the process of remodeling. For key words, we picked *tension, debate, openness, struggling,* and *growing.* For a motto, we chose to quote two statements that Delta educators had made in their discussions with us: "We're having a hard time right now" and "I think we're ready to change."

Beta

Beta was the district that felt most like "school" to many of us on the research team. The district was in the process of talking about moving to a shared decision-making model; however, at the time we were there, most decisions were made in a top-down fashion. Textbooks were adopted by the district, and teachers were expected to use those textbooks.

Decisions that teachers made took place within this framework. As one principal explained,

> Within the parameters of the stated curriculum, within the parameters of the adopted texts, our folks can pretty much make the decisions. [In our school, for example] we don't have one recess time like a lot of the schools do. First grade teachers decide when they are going to have recess. Second grade teachers decide . . . . All they have to do is tell me when that's going to be. That's a kind of decision they make that in some other places they don't.

In all but one of the schools in the district, instruction was centered on the district mandated textbooks. In the school that was an exception, teachers had received permission to use a direct instruction program for both reading and math.

"Assessment" in Beta classrooms predominantly meant tests, and assessment-as-test impacted instruction both directly and indirectly. One teacher explained that teachers tried to cover the curriculum before testing in April:

> We don't think it's right that the kids get tested on what they haven't been taught and so we try to cover everything before the test. It's really a push. Kids get left behind. [Then] after the reading test is over, we can go back and take the time and help the kids that got left behind.

What got taught was also affected by the test, as teachers spent time preparing students for the kinds of things that the tests measured. For example, on one day that we visited, a teacher had written on the board,

> Your principal has said that wearing shorts to school will not be allowed because some students wear them when it's too cold, some wear raggedy cutoffs and it gives students the attitude that they come to school to play instead of learn. Agree or disagree. Explain.

The teacher explained to us that this exercise was part of preparing students to take the holistically scored district writing test and that all year she had been teaching the children to write using a topic sentence, three or four sentences and then a closing sentence.

Tests, then, affected what got taught as well as when it got taught. They also affected how things got taught. On the new (at that time) Illinois State Reading Test, for example, more than one right answer could be correct. One teacher noted that in the past, teachers had taught that there was just one right answer, so that now they had to "change the materials to fit the format [of the state test]."

In a similar fashion, test results had an impact on curriculum at the district level. A central office administrator explained that:

> We found that our students scored very low on antonyms, and what we decided as department was that our goal [would be] to emphasize the instruction of antonyms for our kids.

In Beta, teachers also believed that there was a relationship between tests and textbooks. It was not so much that the district adopted materials to match the tests, but rather that textbook publishers would match their materials to the test. Asked about the influence of the state reading test, for example, one teacher noted, "Tests will change texts and then the teaching will match the text."

Within this framework, accountability seemed to be viewed by teachers as "covering" the district adopted materials. Assessment, at the classroom level, meant assuring that the students had mastered the materials and that the students were prepared for district and state tests. Outside the classroom, administrators noted that the "objective data" provided by these tests were a means of evaluating the quality of the instruction the children received. As one principal explained,

> It's less intuitive and more objective . . . . There is more objective data . . . . Along with this there is the teacher's opinion which is valuable. For example, she might recognize that a child did poorly on the test because he or she was having a bad day . . . . The danger in abandoning the formalized measure is that I often hear teachers' assessments of students and they are incorrect.

We implied above that, by pulling particular characteristics of each district into the foreground, much of the depth and breadth of what we learned would have to be omitted. And part of what was omitted was a discussion of the role of trust throughout the study. We were comfortable not talking about trust in Alpha, Gamma, and Delta, because we were comfortable that trust had been established. Indeed, in all three districts, participants seemed very open and trusting; they seemed comfortable talking about what they perceived to be both the strengths and the weaknesses of their districts, themselves, their classrooms and their students. In Beta, however, it seems important to talk about trust as we were uncomfortable, rather than comfortable, with the issue of trust. In Beta, teachers and principals were "volunteered" to be in the study, and our sense was that if some of them had not been "volunteered," they would not have chosen to participate. Indeed, one of our participants made that point explicitly. Conversations were sometimes strained. Once, a teacher asked to have the tape recorder turned off so that she could talk more freely. Another teacher asked for the interviews not to be taped at all. And, after reading the fieldnotes from the first observation, she commented that she felt uncomfortable having her classroom recorded in such detail. She wondered if it would be possible for the researcher not to take notes during either observations or interviews.

We noted above that we heard voices of individuals in Alpha, voices of teams in Gamma, and voices of debaters in Delta. In Beta, we are not sure what to say about voice. It was certainly a cautious voice, but what did this reflect? Was this a district characteristic? Had we helped to create the sense of distrust? We do not know, and we have no way of knowing. We can report simply that in Beta, we sensed a lack of trust.

In thinking about this issue, we recalled the comments of one administrator, who said to us, "Obviously what happens 99% of the time is that teachers do what they want until caught or whatever—which is standard operating procedure all over the country." Our sense of Beta is best captured visually as a school with closed doors. Our motto for Beta: "Standard operating procedure."

Analysis Across Sites

Having developed a rich portrait of each district, we turned our attention to an analysis of findings across sites. In so doing, we changed the lens used for examination--privileging commonalites over

idiosyncracies—and reversed background and foreground--from trends within districts to districts within trends. When we examined the data across sites, we found that the decision-making model operating within each district exercised a pervasive influence on views and practices for both assessment and instruction. It influenced both the type of instructional decisions as well as the way they were made. It determined the relative value accorded to different types of assessment data (teacher-generated versus test-generated, for example) and even influenced the criteria used by teachers to define informal assessment. Indeed, with regard to the relationship between instruction and assessment, differences among schools and districts were indicative of differences in power relations among administrators and teachers in each district. These themes serve as the bases for the elaboration of our cross-site analysis.

The Impact of Decision-Making Models on Instruction

Instruction in Alpha was characterized by attention to individual children in terms of each teacher's clearly defined vision. One teacher told us that, for her, the essence of teaching was

> touching and changing lives in a positive way allowing students at whatever age to somehow gain a self worth to build that foundation that says, "You're so important and you have so many gifts and so many talents that all we have to do is help you recognize those." And to say to each child that comes into this classroom, "You are a success, the minute you walk in here, you are a success" and to present them with materials and chances and anything that says to them, 'I'm successful.' And once they believe they are successful then you can say, "You're so successful that I want you to take this risk" . . . and it's just a matter of building on the successes and once they're successful and they can feel what that feels like and it feels good then they will be a risk taker. But you can't ask children to come in and be risk takers if they don't know that they are successes. You can change lives but you have to believe that yourself.

Another teacher explained the goals she had for her students in this way:

> I do everything I can to create independence in children which creates self awareness and self evaluation and responsibility for learning.

A principal in one of the Alpha schools responded, when asked what he would do when a conflict arose between a standardized test score and teacher judgement, "I would tend to side with teacher judgment." The superintendent in Alpha epitomized the district attitude in describing his job as trying ". . . to erode school and teacher decision making opportunities as little as possible." Alpha teachers were expected both by administrators and by other teachers to have goals and to make decisions about how to accomplish them.

In Gamma, instruction was characterized by a constant refining of materials and strategies to meet student needs. Teachers were on the lookout for new ideas from workshops, professional reading, and colleagues, and the system supported their exploration. One teacher told us that a strategy we had observed her using had been introduced to her by another teacher who had presented a workshop on an Institute Day. Later she explained how her team had gotten ideas about teaching persuasive writing when they met informally with a consultant who had been brought into the district:

> We had release time and we got to talk with her, my team, for a couple of hours. . . initially, she came to our district and all the fourth grade teachers were released either the morning or afternoon. Then she was brought back and we could request what we wanted her to do for this [second] day. . . we wrote her and said this is what we want you to do and it was talking about persuasive writing.

Gamma administrators told us they had more faith in their teachers than in published materials. In fact, the president of the school board said, "My purpose in visiting a school is to observe and learn. For me to tell a teacher how to teach is inappropriate."

In Beta and Delta, the teacher's job was to cover the curriculum--defined as district-adopted textbooks--and the student's job was to get "it." When teachers did report making instructional decisions, their explanations usually contained references to externally determined materials. When one teacher in Delta explained the flexibility she had to change students' placements in the reading program, she said:

> Basically, that's what they ask of us--that we make sure that we test them and make sure that they've mastered the basic skills that are required. . . .The only thing that would ever come back to me would be if I didn't finish reading or math...In first grade you have to [cover it all] and there [are] no questions.

Even the purposes given to students for what they were learning reflected this focus on adopted materials. For example, we observed a third-grade teacher who introduced the *are/air* spelling patterns to her class by explaining, "You need to know this today for one of your workbook pages."

Administrators in these two districts, Beta and Delta, explained that they monitored instruction to be sure that teachers were covering materials in the textbooks at an adequate pace. As one Delta administrator told us, the tests "give the principal a handle on whether they [the teachers] are moving along through the levels they're expected to cover."

Additionally, in Delta, the materials budget had been removed from the schools and it was the district that selected workbooks and other ancillary materials used in classrooms. Supplementing with the Weekly Reader was described by the superintendent as an "approved exception."

If we examine instruction through the filter of decision making we can better understand why two classrooms can use the same tools--for example, the same text materials and the same tests--and yet be very different places. In one classroom, the materials would be there because the teacher, informed by her goals, her knowledge, and her students, had decided they were the best to use. But in the other, the materials would be there because someone outside the classroom decided to mandate their use and to monitor the teacher's compliance with the mandate. Our data suggest that these two classrooms would turn out to be very different places in which to be a teacher and to be a learner. Implicit, of course, in our line of reasoning is the conclusion that decision making is a process in which the distribution of power and authority is a central, perhaps the central, issue.

The Impact of Decision-Making Models on Assessment

We saw similar patterns emerge when we looked at the relationship between assessment and the decision making model of each district. The differences did not manifest themselves in the particular types of assessment used: all four districts gave nationally normed standardized tests, administered the Illinois State Reading test, used informal observations, and relied on daily work samples to assess their students' progress. Yet these assessments played out differently across districts: There were important differences in the relative value different districts accorded to assessment data from tests and to assessment data from teachers.

In Alpha and Gamma, tests were viewed as inconveniences to be dealt with as efficiently as possible in order to get on with the business of schooling. The superintendent in Alpha reported that he had reservations about standardized tests and was concerned that they took time away from instruction. Alpha principals told us that they got no useful information from tests, and that when a conflict arose between test data and teacher judgment, they valued the teacher's judgment more highly. A central

office administrator in Gamma told us that standardized test scores were meaningless because of the poor quality of the tests, and explained "It's so much easier to assess things that aren't important. You're better off doing no assessment than to assess the trivial things." When asked if she considered teacher judgment comparable to formal assessment measures, one Gamma principal answered, "Maybe even more so."

In Beta and Delta, test data were highly valued in part because they provided a means to avoid relying on teacher judgment. Teachers were required to use end-of-unit and end-of-level basal tests in Delta; in fact, one administrator reported that he did not know how teachers would know what to do for individual students without the computerized prescriptions that went along with these tests. Even when unsure about the quality of tests, Delta administrators preferred them to teacher judgment. A central office administrator told us, "whether it's really good or not, that's immaterial. It's some kind of screening device. Then the teacher . . . will know approximately where a child stands." Even when administrators were critical of the basal tests, they implicitly acknowledged that they believed the tests to be superior to teacher judgment. As one principal said,

> The tests are embarrassing, the unit tests are embarrassing. They rarely focus on instruction, comprehension. They're about isolated skills. It doesn't give you a full picture. [But] we do track students with [them] because that's all we have.

In Beta, principals were required to meet with their staff each year to review the performance of their students on the SRA, and were expected to use this information to "sharpen their focus" and to determine areas of emphasis for the following year. Because Beta educators wanted to ensure that the students were ready for the next level, the goal of these meetings was to identify, for teachers, the areas of importance on which they should concentrate their efforts. Otherwise, as one central office administrator told us, "if we have a curriculum and we just get to do all kinds of fun little things, but no purpose, no goal, I don't think we are being fair to teachers or students."

One finding surprised us all. We found little relationship between the assessment fervor within a district and its actual ranking on external accountability indices. For example, Delta officials (not teachers, but administrators) volunteered to be a part of our study because they wanted to better understand the causes for their low scores on the state reading test. Their pervasive concern about assessment stands in stark contrast with the almost cavalier attitude toward external accountability indices taken by all the professional staff (teachers and administrators) in Alpha. Based upon these attitudes, one might have predicted that Alpha was near the top of the distribution of state scores, while Delta was nearer the bottom; however, the data in Table 1 suggest just the opposite. Perhaps the combination of a top-down decision-making structure, which allows administrators to define problems as well as select solutions, and its location near other high-scoring suburban districts, such as Gamma for example, colored Delta officials' perception about the nature and severity of their district reading problem. Perceptions about the meaning of assessment data, like the data themselves, are highly situated.

The Impact of Decision-Making Models on Informal Assessment

We can further our understanding of the relationships among assessment, instruction, and decision making by looking at the criteria used by teachers for informal, classroom assessment of student progress. In Alpha and Gamma, the criteria teachers used for informal assessment most often originated in their own goals and philosophies, and if a student were found to be having trouble, this tended to be attributed to a wide variety of possible explanations. One Alpha teacher explained how she made sense of the information she gathered informally:

> What I try to be is what the Quakers call "mindful" of what they're doing and I try to analyze what I'm seeing. I try to understand what the information I'm getting is actually telling me.

A fourth-grade teacher in Gamma reported that informal observation could give her feedback on her teaching. She said:

> I guess I just walk around and see if they're doing what they're supposed to be doing and if they're not then I might ask them if they understood what they were supposed to be doing and if I find that [several] kids aren't understanding then I'll explain and clarify [my instructions].

An Alpha teacher reported that she had individual goals for students, and that she tried to "read" her students to determine how to respond to them in terms of these goals:

> I think so much of what I do is intuitive now. I read body English pretty well. I read faces pretty well. I make it safe in here for them to ask questions and to share their ideas. They really draw out of me what it is that they need to know. One of the things I do is I try to understand the kind of information that they're seeking.

Another Alpha teacher told us that she observed students in as many different situations as possible to inform her instructional decisions:

> We have a period of having the children learn how to get along in this room . . . and just seeing what kind of learners they are--what are their habits? . . . We do a lot of observations. We throw them into a lot of interacting activities . . . . I find there's nothing that matches just having every occasion possible to talk one to one, to look at work daily and weekly . . . . Evaluation is daily--all the time. It isn't just the end of a period type of thing.

In contrast, when teachers in Beta and Delta talked about their own informal observations and daily classroom work, they frequently referred to the externally determined curriculum and their concern that students "get it." One Beta teacher explained her system for recording informal observations:

> After a while you really know them, you know which ones can do everything. In workbooks, I try to keep track of vocabulary and comprehension, and then on my SRA, I write down the number right and the number wrong.

Teachers in Delta also relied heavily on workbook pages for information about student learning. As one teacher said,

> I do occasionally, when I find on the workbook that there is some confusion, I will go to the skill pack and take out those pages and staple them together and re-teach.

A Beta teacher also described how her observations of daily work informed later teaching:

> I try, if there is something that really stands out, I'll jot a note to myself. For example, the other day with the red reading group, I noticed that Jane was having some difficulty with the vocabulary words. When we were doing it together she was the one who was not saying it correctly, so I could pull her aside later and work with her and mention to the Chapter I teacher that she may need extra help with that lesson.

In Beta and Delta, several of the teachers told us that if students consistently demonstrated that they were having trouble getting "it," if they were having trouble with understanding the task, that would indicate to them that the students were incorrectly placed in the material and they would move them to a lower level. As one Delta teacher explained,

> If I have someone who is really having difficulty, I have gone through the remedial steps, I have done the reteaching and given them additional work, and they still seem to be having a problem--and by seem to be having a problem, I mean the worksheets they are doing for me, discussion, writing they are doing for me, I can see that the same problem is showing up that is supposedly corrected--if this continues then we would change the placement immediately rather than wait until next year.

Conclusions and Reflections

When we began this study, we were seeking to understand the relationship between assessment and instruction. We situated our study in the context of decision making in the classroom, in the building, and in the district. In this way, we hoped to ferret out the subtle influences that assessment might have on instruction. However, what we found was that the salient relationship was not between assessment and instruction per se. Granted, the two were related, but their relationship was moderated by the decision-making model of the district.

Generally, when assessment-as-test did appear to "drive" instruction, this relationship seemed to be an artifact of a model in which the people responsible for delivering instruction had little authority and power. Teachers were responsible for instruction but administrators had instructional power; similarly, central office staff controlled principals' building-level decision making, and publishers and state agencies strongly affected the decision making of central office staff. All along the way, members of the system were accountable to external forces.

When assessment-as-test did *not* appear to "drive" instruction, the controlling decision-making model was one in which individuals maintained the authority to make decisions according to their individual and collective philosophies. These decisions were characterized by responsibility to individual learners rather than by accountability to outside sources.

Two Hypotheses

These understandings lead us to two major hypotheses, both of which raised for us concerns and new questions.

First, assessment-as-test may not always drive instruction. Indeed, based on our research, we believe that it is not tenable to talk about the relationship of assessment to instruction. The relationships among assessment, instruction, and decision making seem to be much more complex than originally thought, and the complexity extends across many dimensions: they are complex culturally, socially, politically, and historically. This hypothesis of unexpected complexity, grounded in the data from these four sites, provides a challenge to the causal, linear notions of assessment and instruction that now dominate the literature. To understand assessment and instruction, we believe that it is first necessary to understand the complexities of "doing school" at any given location.

Second, when assessment-as-test does drive instruction, it may not take us where we want to go. In districts in which we could trace a direct relationship from scores on standardized measures to classroom practices, what we observed could hardly be called admirable. In one district, students were subjected to a daily 10-minute test blitz of skills to prepare them for a state test. In another, teachers rushed to finish the year's work before testing in April. In yet another, we saw the inclusion of topical-knowledge

measures on the Illinois State Reading Test translated into a requirement that students as low as Grade 1 complete prior-knowledge worksheets before every basal story.

Broader Concerns

Teacher prerogative. We were troubled by the lack of teacher voice in the assessment-driven districts. Teachers in Alpha and Gamma talked about people whose work they had read, heard at conferences, and "thought with," including each other. For the most part, this did not happen in Beta and Delta. Indeed, the pattern seemed to be that in districts with higher levels of autonomy, teachers conveyed philosophies and visions of what they wanted to accomplish. In districts with less teacher autonomy, teachers talked predominantly (although not necessarily positively) about what others wanted them to accomplish, echoing Smith's (1991) description of "the teacher after testing reform:"

> Far from the reflective practitioner or the empowered teacher, those optimistic images of the 1980s, the image we project of teachers in the world after testing reform is that of interchangeable technicians receiving the standard curriculum from above, transmitting it as given (the presentation manual never leaving the crook of their arms), and correcting-multiple choice responses of their pupils. (p. 11)

These practices, practices that we observed in particular schools in particular districts, are not isolated events. In fact, they are consistent with findings that have been observed elsewhere. For example, based on data collected over 15 months in two schools in the Phoenix metropolitan area (Smith, Edelsky, Draper, Rottenberg, & Cherland, 1989), Smith (1991) reported that external testing had a number of negative effects on teachers and teaching, including a tendency to use instructional methods that resemble testing:

> Take away the publishers' trappings, and one would be hard pressed to distinguish an ITBS item from a question on a typical worksheet. Both call for the pupil to select among alternative options the one that an outside expert has decided in advance is correct. (p. 10)

Shepard and Dougherty (1991) suggested that "improvement," as measured by standardized tests, may come at the expense of reduced instructional time, increased stress, and demoralizing effects on teachers and students.

If our findings, along with those reported by others, accurately characterize the impact of assessment on teacher voice and teacher prerogative, then the use of assessment as a lynchpin in the current educational reform movement creates a dilemma for those who support the current movement toward the systemic reform of schools. If one tenet of the reform movement is the greater empowerment of teachers in local decision making and another is reliance on new and more responsible assessments for accountability, a collision of forces may be inevitable. Assessments, particularly those that are externally imposed, may be inherently incapable of privileging individual teacher voice and prerogative.

Rethinking how to study assessment and decision making. If we were starting this study today, we would want to broader the scope demographically and methodologically. We are all too aware of the situated character of our findings and insights. We set our study in Illinois, partly out of convenience (we were located there) and partly because of the notoriety of the assessment-driven state reform plan that was being implemented as we began the study. It would be informative to study the same sets of decision-making relationships in schools and states that have, for example, formally committed themselves to perfomance and portfolio assessment systems, both for classroom decision making and accountability reporting. Methodologically, we chose to do "snapshots" of key players (educators and students) within classrooms within schools within districts. In another trip across this landscape, we

would likely use the metaphor of "portrait over time" to guide our search, perhaps studying fewer sites in more detail and for longer periods of time. And more than likely, we would develop a more situated view of concepts and categories, opening ourselves to the possibility that all of them might not extend across sites.

Rethinking school-based research and development. Because we now understand that school and district cultures impact the relationship between assessment and instruction in particular classrooms and particular schools, we realize that the complexities of schools as cultures demand complex, not simple, agendas for change and that research on schools needs to consider the broader political contexts operating within schools and districts. Assessment and instruction can no longer remain isolated concepts; their understanding requires a wider, cultural lens. Research questions can no longer be framed as finite—limited to the classroom, teacher, or students. Instead, the wider school environment, and the relationships within, must be explored, understood, and used to inform any change agenda. For example, as teachers try to maximize learning for all students by appropriately linking assessment and instruction, what, if anything, within the school culture would need to change? Who would determine the change agenda? How would the new agenda be implemented? What roles would teachers, administrators, community members, parents, and students assume?

Having acknowledged the cultural differences and the political complexities of "doing school," we also need to ask how one district's experience can inform another, how outside sources can provide assistance, and how all schools can distinctly but consistently move positively forward. Our hope is that the data from this study will advance a cultural perspective on change and decision making, inform the change agendas that educators across diverse cultures set for themselves, and contribute to understanding some of the questions various stakeholders pursue.

References

Berlak, H. (1978). Testing in a democracy. *Educational Leadership, 35*, 17.

Brandt, R. (1978). The search for solutions. *Educational Leadership, 35*, 3.

Brookover, W. B. (1987). Distortion and overgeneralization are no substitutes for sound research. *Phi Delta Kappan, 69*, 225-227.

Burry, J. (1981). *The design of testing programs with multiple and complementary uses* (Report Number 165). Los Angeles: UCLA Center for the Study of Evaluation.

Calderhead, J. (1988). *Reflective teaching and teacher education*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Calfee, R. (1987). The school as a context for the assessment of literacy. *The Reading Teacher, 40*, 738-743.

Center for the Study of Testing, Evaluation, and Educational Policy. (1992, October). *The influence of testing on teaching math and science in grades 4-12* (Vol. 1-5). Boston: Boston College, Center for the Study of Testing, Evaluation, and Educational Policy.

Cohen, M. (1988). Designing state assessment systems. *Phi Delta Kappan, 70*, 583-588.

Conner, K., Hairston, J., Hill, I., Kopple, H., Marshall, J., Scholnick, K., & Schulman, M. (1985). Using formative testing at the classroom, school and district levels. *Educational Leadership, 43*, 63-68.

Dorr-Bremme, D. W., & Herman, J. L. (1986). *Assessing student achievement: A profile of classroom practices*. Los Angeles: University of California, Center for the Study of Evaluation.

García, G. E., & Pearson, P. D. (1994). In L. Darling-Hammond, *Review of research in education* (Vol. 10). Washington, DC: American Educational Research Association.

García, G. E., Pearson, P. D., & Jiménez, R. (1994). *The at-risk situation: A synthesis of reading research*. Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Chicago: Aldine.

Glesne, C., & Peskhin, A. (1992). *Becoming qualitative researchers*. New York: Longman.

Graves, D., & Sunstein, B. (1992). *Portfolio portraits*. Portsmouth, NH: Heinemann.

Guthrie, J., & Lissitz, R. (1985). A framework for assessment-based decision making in education. *Educational Measurement: Issues and Practices*, 26-30.

Haladyna, T. M., Nolen, S. B., & Haas, N. S. (1991). Raising achievement test scores and the origins of test score pollution. *Educational Researcher, 20*, 2-7.

Hancock, J., Turbill, J., & Cambourne, B. Assessment and evaluation of literacy learning. In S. W. Valencia, E. H. Hiebert, & P. P. Afflerbach, (Eds.), *Authentic reading assessment* (pp. 46-62). Newark, DE: International Reading Association.

Haney, W. (1984). Testing reasoning and reasoning about testing. *Review of Educational Research, 54*, 597-654.

Haney, W. (1985). Making testing more educational. *Educational Leadership, 43*, 4-13.

Hansen, J. (1992). Literacy portfolios: Helping students know themselves. *Educational Leadership, 49*, 66-68.

Hansen, J. (1994). Literacy portfolios: Window on potential. In S. W. Valencia, E. H. Hiebert, & P. P. Afflerbach (Eds.), *Authentic reading assessment* (pp. 26-40). Newark, DE: International Reading Association.

Herman, J. L., & Golan, S. (n. d.). *The effects of standardized testing on teaching in schools*. Los Angeles: The National Center for Research on Evaluation, Standards, and Student Testing, UCLA Graduate School of Education.

Johnston, P. (1987). Steps towards a more naturalistic approach to the assessment of the reading process. In J. Algina (Ed.), *Advances in context based educational assessment*. Norwood, NJ: Ablex.

Koretz, D. M., Linn, R. L., Dunbar, S. B., & Shepard, L. A. (1991, April). *The effects of high-stakes testing on achievement: Preliminary findings about generalization across tests*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

Madaus, G. F. (1985). Test scores as administrative mechanisms in educational policy. *Phi Delta Kappan, 55*, 611-617.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.

Pearson, P. D. (1992). Reading. *The encyclopedia of educational research*, 1075-1085.

Pearson, P. D., & Valencia, S. (1987). Assessment, accountability, and professional prerogative. In J. E. Readence & R. S. Baldwin (Eds.), *Research in literacy: Merging perspectives: Thirty-sixth yearbook of the National Reading Conference* (pp. 3-16). Rochester, NY: National Reading Conference.

Popham, W. J., Cruse, K. L., Rankin, S. C., Sandifer, P. D., & Williams, P. L. (1985). Measurement driven instruction: It's on the road. *Phi Delta Kappan, 55*, 628-634.

Rodriguez, A., Stephens, D., Commeyras, M., Stallman, A. C., Shelton, J., Pearson, P. D., Roe, M., Weinzierl, J., & Gilrane, C. P. (1993). *Assessment and decision making in Delta* (Tech. Rep. No. 591). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Rothman, R. (1992, October 21). Study confirms "fears" regarding commercial tests. *Education Week, 12*, 1, 13.

Shelton, J., Stephens, D., Stallman, A. C., Commeyras, M., Pearson, P. D., Roe, M., Rodriguez, A., Kondrot, J., & Weinzierl, J. (1993). *Assessment and decision making in Gamma* (Tech. Rep. No. 592). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Shepard, L. A. (1989). Why we need better assessments. *Educational Leadership, 46,* 4-9.

Shephard, L. (1990). Inflated test score gains: Is the problem old norms or teaching to the test? *Educational Measurement: Issues and Practices, 20,* 15-22.

Shepard, L. A., & Dougherty, K. C. (1991). *Effects of high-stakes testing on instruction.* Paper presented at the annual meeting of the American Educational Research Association and the National Council on Measurement in Education, Chicago.

Simmons, W., & Resnick, L. Assessment as the catalyst of school reform. *Educational Leadership, 50,* (5), 11-16.

Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher, 20,* 8-11.

Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1989). *The role of testing in elementary schools.* Los Angeles: UCLA, Center for Research on Educational Standards and Student Tests, Graduate School of Education.

Stedman, L. C. (1987). It's time we changed the effective schools formula. *Phi Delta Kappan, 69,* 215-224.

Stephens, D., Pearson, P. D., Stallman, A. C., Shelton, J., Commeyras, M., Roe, M., Rodriguez, A., & Moll, J. (1993). *Assessment and decision making in Alpha* (Tech. Rep. No. 589). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Tierney, R. J., Carter, M. A., & Desai, L. E. (1991). *Portfolio assessment in the reading-writing classroom.* Norwood, MA: Christopher-Gordon.

Tierney, R. J., & McGinley, W. (1987). *Serious flaws in written literacy assessment.* Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Valencia, S., McGinley, W., & Pearson, P.D. (1990). Assessing reading and writing. In G. G. Duffy (Ed), *Reading in the middle school* (2nd ed., pp. 124-153). Newark, DE: International Reading Association.

Valencia, S., & Pearson, P. D. (1987). Reading assessment: Time for a change. *Reading Teacher, 40,* 726-732.

Weinzierl, J., Stephens, D., Stallman, A. C., Pearson, P. D., Shelton, J., Rodriguez, A., Roe, M., Commeyras, M., Clark, C., & Moll, J. (1993). *Assessment and decision making in Beta* (Tech. Rep. No. 590). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Wixson, K. K., Peters, C. W., Weber, E. M., & Roeber, E. D. (1987). New directions in statewide reading assessment. *The Reading Teacher, 40,* 749-754.

Author Note

The research described in this report was sponsored by funds from the Reading Research and Education Center from the Office of Educational Research and Improvement.

This was a long, complex study, and different players contributed in different proportion at different points in its conduct. For example, Anne Stallman and Michelle Commeyras played the major roles in the early data collection and analyses and were supported by Mary Roe and Alicia Rodriguez. Judy Shelton, Alicia Rodriguez, and Janelle Weinzerl were heavily involved in analyzing data for and writing the individual case studies; Anne Stallman and Michelle Commeyras supported them in this effort. Colleen Gilrane conducted the analysis for the cross-site analysis and created an initial draft of the current manuscript. Mary Roe played a major role in the revision process that led to the current version. Diane Stephens and P. David Pearson were involved in every phase of the work.

Table 1

Demographic Information for Participating Districts from State of Illinois Report Card

	Total Enrolled	Expend per pupil	% attend rate	% low income	% mobility rate	% not promoted	% white enrol	% Afr Amer enrol	% Latino enrol	% Asian Amer enrol	% Native Amer enrol	% LEP enrol	Average Gr 3 State Rdg Test
Alpha	4990	3999	94	33	24	2	71	20	7	7	.1	6	254
Alpha 1	542		96	26	17	1	80	16	2	2	0	2	273
Alpha 2	552		96	24	13	2	86	12	2	2	0	.7	261
Beta	8235	3271	94	21	23	2.4	67	29	.9	3	.1	.4	254
Beta 1	453		94	33	35	2.5	62	34	.4	3	0	0	229
Beta 2	451		95	28	21	4.4	65	30	.9	4	0	0	270
Delta	3201	3552	94	7	21	3	82	.5	12	5	.1	9	264
Delta 1	416		94	1	13	1	95	.5	3	2	0	1	252
Delta 2	298		95	5	21	.3	85	1	8	6	0	4	257
Gamma	15191	3434	95	.5	12	.7	93	1	.6	5	.2	.5	302
Gamma 1	463		95	.2	33	.4	94	1	1	3	.4	.2	272
Gamma 2	693		96	.1	13	.1	92	1	.6	7	0	.4	308
State		4008	93	29	21	3	67	22	9	2	.1		250

Appendix A:

Interview Questions

Teacher

I would like to ask you a number of questions regarding the role of assessment and the decision-making process in your district. Of course, your comments will be considered confidential and we will not identify your opinions by name or school district.

1. Please give me a picture of the decision-making process in your district.

   a. Please give me a hypothetical situation so I might better understand how this works.

2. What kinds of decisions do you make as a teacher?

   a. Is there anything done or expected that extends or limits this decision-making?

   b. What is the general policy about curriculum? Who formulates it? What effect does this policy have on classroom decision making?

   c. Who chooses the materials that are used in the classroom? How and to what degree do those materials influence decision making in the classroom?

   d. What is the relationship between tests, material selection, instructional strategies and instructional decisions?

3. How is student progress monitored in your classroom?

   a. What sorts of formal and informal assessment take place?

   b. How is the data used?

   c. How does your monitoring of student progress, assessment, and data usage compare to other teachers?

4. How do you think people in your district feel about the decision-making process? The assessment process? What do you think are the prospects for change in either of these areas?

Appendix A (Cont.)

Interview Questions

Superintendent

I would like to ask you a number of questions regarding the role of assessment and the decision-making process in your district. Of course, your comments will be considered confidential and we will not identify your opinions by name or school district.

1. Please give me a picture of the decision-making process in your district.

   a. Please give me a hypothetical situation so I might better understand how this works.

   b. (Make sure you know the types of decisions the superintendent makes and those he or she considers the responsibility of principals, teachers, school board members, or other personnel.)

2. What kinds of decisions do you expect teachers to make?

   a. Is there anything done or expected that extends or limits this decision making? or

      Is there anything done or expected by the administration that extends or limits this decision making?

   b. What is the general policy about curriculum? Who formulates it? What effect does this policy have on classroom decision making?

   c. Who chooses the materials that are used in the classroom? How and to what degree do you expect those materials to influence decision making in the classroom?

   d. What do you think the relationship should be between tests, material selection, instructional strategies and instructional decisions?

3. How is student progress monitored in your district?

   a. Do you expect this to be the same from building to building?

   b. What sorts of formal and informal assessment take place? Is this similar from building to building?

   c. How is the data used?

4. How do you think people in your district feel about the decision-making process? The assessment process? What do you think are the prospects for change in either of these areas?

Appendix B

Observation and Interview Coding Systems

Interview codes

Slot 1
Talking about
a. Self
b. Superintendent
c. Assistant superintendent
d. Board member
e. Staff development (person)
f. Consultant
g. Principal
h. Teacher
i. Student
j. Parent
k. State
l. District
m. Administration
n. Staff development (program)
o. Decision making
p. Curriculum
q. Instruction
r. Assessment
s. Discipline
t. Materials
u. Classroom
v. School
w. Committees
x. Town
y. PTA
aa. Assistant principal
ab. Social worker
ac. Education
ad. Budget
af. Salesperson

Slot 2
Type
21 Philosophy
22 Policy
23 Practice

Slot 3
Source
301 Mandate
302 Board of education
303 Superintendent
304 Principal
305 Colleague
306 Staff development
307 Book
308 Teacher education
309 Personal experience
310 Experience as a student
311 Teaching experience
312 Intuition
313 Can't identify
314 State
315 Professional meeting

Slot 3 continued
316 Reflection
317 Source
318 Assistant superintendent

Slot 4
Control
401 Self
402 Cooperative
403 Committee
404 Teacher
405 Principal
406 School
407 District
408 State
409 Student
410 Aide
411 Superintendent
412 Assistant superintendent
413 PTA
414 Union
415 Other administrator
416 Board
417 Parent

Slot 5
Type of participation
51 Mandatory
52 Voluntary

Slot 6
Type of assessment
601 National
602 State
603 District
604 Text publisher
605 Other publisher
606 Teacher made
607 Samples
608 Checklist
609 Informal
610 Dynamic

Slot 7
Uses/role of assessment
701 Accountability
702 Program evaluation
703 Teacher evaluation
704 Pupil placement
705 Reporting pupil progress
706 Monitoring pupil progress
707 Choosing materials
708 Instructional decisions
709 Diagnosis

Observation codes

Slot 1
Task definition
1 Assessment
2 Behavior management
3 Classroom activity
4 Planning/schedule
5 Nonacademic

Slot 2
Grouping
201 Whole/T
202 Small/T
203 Individual/T
204 Whole/NoT
205 Small/NoT
206 Individual/NoT

Slot 3
Content
301 Social studies
302 Science
303 Math
304 Literature
305 Reading
306 Writing
307 Grammar
308 Spelling
309 Phonics
310 Vocabulary
311 Music
312 P.E.
313 Drama
314 Art
315 Other
316 Health

Slot 4
Materials
400 None
401 Text
402 Basal
403 Trade book
404 Workbook/worksheet
405 Blank paper
406 Kit
407 Manipulative
408 Computer
409 Tape recorder
410 Other gadgets
411 Art supplies
412 Chalkboard
413 Homemade book
414 Reference material
415 Test
416 Other
417 Film/movie

Slot 5
Type of activity
(Use only with Slot 1 #3)
501 Telling
502 Teacher initiates/ student responds
503 Scaffold
504 Discussion

Slot 6
Type of assessment
(Use only with Slot 1 #1)
601 National
602 State
603 District
604 Text publisher
605 Other publisher
606 Teacher made
607 Samplers
608 Checklist
609 Informal
610 Dynamic

BEST COPY AVAILABLE